THE ASTRONOMICAL JOURNAL, 122:1163–1171, 2001 September

PHOTOMETRIC REDSHIFTS FROM RECONSTRUCTED QUASAR TEMPLATES

TAMÁS BUDAVÁRI,[1,2] ISTVÁN CSABAI,[1,2] ALEXANDER S. SZALAY,[2] ANDREW J. CONNOLLY,[3] GYULA P. SZOKOLY,[4] DANIEL E. VANDEN BERK,[5] GORDON T. RICHARDS,[6] MICHAEL A. WEINSTEIN,[6] DONALD P. SCHNEIDER,[6] NARCISO BENÍTEZ,[2] J. BRINKMANN,[7] ROBERT BRUNNER,[8] PATRICK B. HALL,[9,10] G. S. HENNESSY,[11] ŽELJKO IVEZIĆ,[9] PÉTER Z. KUNSZT,[2] JEFFREY A. MUNN,[11] R. C. NICHOL,[12] JEFFREY R. PIER,[11] AND DONALD G. YORK[13]

Received 2001 April 3; accepted 2001 June 8

ABSTRACT

From Sloan Digital Sky Survey (SDSS) commissioning photometric and spectroscopic data, we investigate the utility of photometric redshift techniques in the task of estimating QSO redshifts. We consider empirical methods (e.g., nearest neighbor searches and polynomial fitting), standard spectral template fitting, and hybrid approaches (i.e., training spectral templates from spectroscopic and photometric observations of QSOs). We find that in all cases, because of the presence of strong emission lines within the QSO spectra, the nearest neighbor and template-fitting methods are superior to the polynomial-fitting approach. Applying a novel reconstruction technique, we can, from the SDSS multicolor photometry, reconstruct a statistical representation of the underlying SEDs of the SDSS QSOs. Although the reconstructed templates are based on only broadband photometry, the common emission lines present within the QSO spectra can be recovered in the resulting spectral energy distributions. The technique should be useful in searching for spectral differences among QSOs at a given redshift, in searching for spectral evolution of QSOs, in comparing photometric redshifts for objects beyond the SDSS spectroscopic sample with those in the well-calibrated photometric redshifts for objects brighter than 20th magnitude, and in searching for systematic and time-variable effects in the SDSS broadband photometric and spectral photometric calibrations.

Key words: galaxies: distances and redshifts — galaxies: photometry — quasars: general — methods: statistical

1. INTRODUCTION

In recent years, with new multicolor surveys coming online, the application of photometric redshifts to the analysis of the physical properties of galaxies has become increasingly popular (see Weymann et al. 1999 for recent works). Photometric redshift estimation techniques rely on detecting the passage of continuum features within the spectral energy distribution (SED) of sources (e.g., for galaxies, the 4000 Å break) across a series of photometric passbands. For objects with featureless spectra, photometric redshift estimation is extremely difficult or impossible, as the colors of these objects vary slowly (if at all) with redshift. Emission lines may, however, help if they are strong enough relative to the continuum flux to be detectable within the broadband photometry. As quasar emission lines carry a significant amount of flux (Francis et al. 1991; Richards et al. 2001a), we expect that they will provide sufficient leverage to help in the redshift prediction for lower redshift objects. In addition, it may be easier to estimate redshifts for high-redshift objects ($z \gtrsim 3$), as a result of the blanketing of the Lyα forest (see, e.g., Madau 1995). In the astronomical literature, there are a number of different approaches for estimating photometric redshifts. Most techniques can be categorized into two basic classes: those that use SEDs as spectral templates derived from models or observations (Koo 1985; Gwyn & Hartwick 1996; Sawicki, Lin, & Yee 1997; Connolly et al. 1999; Fernández-Soto, Lanzetta, & Yahil 1999; Benítez 2000; Budavári et al. 1999, 2000; Csabai et al. 2000) and those that establish a direct empirical relation between colors and redshift (Connolly et al. 1995; Wang, Bahcall, & Turner 1998; Brunner, Connolly, & Szalay 1999) using a training set.

In this paper and the companion paper by Richards et al. (2001b), we investigate these techniques and determine how successfully they may be applied to the question of the estimation of the redshifts of QSOs. In this study, we put more emphasis on the so-called template-fitting method and quasar template reconstruction. In the companion paper, we discuss an empirical method that essentially assumes that quasar colors are alike at a given redshift. We use a set of $\sim$2600 known QSOs with five-band ($u'g'r'i'z'$; Fukugita et al. 1996; Lupton et al. 2001) Sloan Digital Sky Survey (SDSS; York et al. 2000) photometry and redshifts in the $0 < z < 5$ range. In Richards et al. (2001b), we study the effects caused by reddened quasars and extended objects

[1] Department of Physics of Complex Systems, Eötvös Loránd University, Pf. 32, H-1518 Budapest, Hungary.
[2] Department of Physics and Astronomy, Johns Hopkins University, 3400 North Charles Street, Baltimore, MD 21218-2686.
[3] Department of Physics and Astronomy, University of Pittsburgh, 3941 O'Hara Street, Pittsburgh, PA 15260.
[4] Astrophysikalisches Institut Potsdam, An der Sternwarte 16, D-14482 Potsdam, Germany.
[5] Fermi National Accelerator Laboratory, P.O. Box 500, Batavia, IL 60510.
[6] Department of Astronomy and Astrophysics, 525 Davey Laboratory, Pennsylvania State University, University Park, PA 16802.
[7] Apache Point Observatory, P.O. Box 59, Sunspot, NM 88349.
[8] Department of Astronomy, 105-24, California Institute of Technology, 1201 East California Boulevard, Pasadena, CA 91125.
[9] Princeton University Observatory, Peyton Hall, Princeton, NJ 08544-1001.
[10] Departamento de Astronomía y Astrofísica, Pontificia Universidad Católica de Chile, Casilla 306, Santiago 22, Chile.
[11] US Naval Observatory, 3450 Massachusetts Avenue, NW, Washington, DC 20392-5420.
[12] Department of Physics, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15232.
[13] Deparment of Astronomy and Astrophysics, University of Chicago, 5640 South Ellis Avenue, Chicago, IL 60637.

concerning redshift outliers and also present more examples of science that will benefit from QSO photometric redshifts. Observational and target selection details are also given there.

In § 2, this study considers empirical methods to estimate redshift. In § 3, we apply the standard template-fitting technique using composite spectra, and in §§ 4 and 5 we extend our analysis to include SEDs reconstructed from photometric information.

2. EMPIRICAL APPROACH

The simplest way of getting photometric redshifts is empirical fitting. We assume the redshift, z, to be a simple function of the magnitudes or colors. Connolly et al. (1995) used low-order polynomial functions for galaxies. The fitting formula (redshift vs. colors) is derived by using a training set of photometric data *and* spectroscopic redshifts. After the calibration, this simple analytic function can predict redshifts for objects directly from photometric observations. While this approach has proved successful for studies of the photometric redshifts of galaxies, the SEDs of quasars differ from those of galaxies in many ways. In the optical and UV, a quasar spectrum consists of an approximately power-law continuum together with strong emission and absorption lines. Thus, quasar colors can change drastically over the space of a small interval in redshift (as an emission line passes from one passband to the next) or can remain constant with redshift as a result of the power-law dependency of the continuum. Given this situation, we would not expect polynomial fitting to yield accurate and reliable photometric redshifts for QSOs, because of the fact that the polynomial functions used vary slowly with redshift. This preconception is borne out by the SDSS data. Fitting both quadratic and cubic relations between the colors of the quasars and their redshifts yielded almost no correlation between prediction and spectroscopic redshifts.

A more adaptive technique than polynomial fitting is that of nearest neighbor (NN) mapping. This technique assigns the redshift of a source to that of the closest object in color space (with a known redshift). This approach is, essentially, a piecewise constant function. Although, as we stated above, the polynomial technique does not work, a combination of the polynomial technique and the NN fitting technique can also be applied in which we find the nearest neighbors and locally fit the polynomial relation over objects within a given radius. This method can suppress any unwanted effect due to possible erroneous nearest neighbors, but it would require more data. A similar empirical method is developed in the companion paper (Richards et al. 2001b), which implements an NN estimator where the reference points are derived from colors averaged over redshift bins.

Figure 1 shows the correlation between the photometric and spectroscopic redshifts from applying the NN estimator. Most of the predicted redshifts ($\sim$70%) match well the spectroscopic measurements. There are, however, a significant number of outliers (i.e., the distribution of the prediction error is clearly non-Gaussian). This is due to the color degeneracy, i.e., within the photometric uncertainties, identical objects (in colors $u'-g'$, $g'-r'$, $r'-i'$, $i'-z'$) have significantly different spectroscopic redshifts. For this reason, for this figure and for all of the following spectroscopic redshift ($z_{\rm spec}$) versus photometric redshift ($z_{\rm phot}$) comparison figures, we calculate two rms values: one for all of the objects ($\Delta_{\rm all}$) and one for the objects that have $|z_{\rm spec} - z_{\rm phot}| < 0.3$ ($\Delta_{0.3}$). The former of these measures is used to estimate the effect of the contamination of a QSO photometric redshift sample by catastrophic failures, and the latter statistic is to measure the intrinsic accuracy of the photometric redshift relation for different analysis techniques. For the NN estimator these values are $\Delta_{\rm all} = 0.64$ and $\Delta_{0.3} = 0.116$. Because of the degeneracy, the number of outliers ($\Delta z > 0.3$) is quite large, approximately 30%. We also note that these results are for a sample of objects already known to be quasars. If the input sample includes objects that turn out not to be quasars, then our results will be diluted.

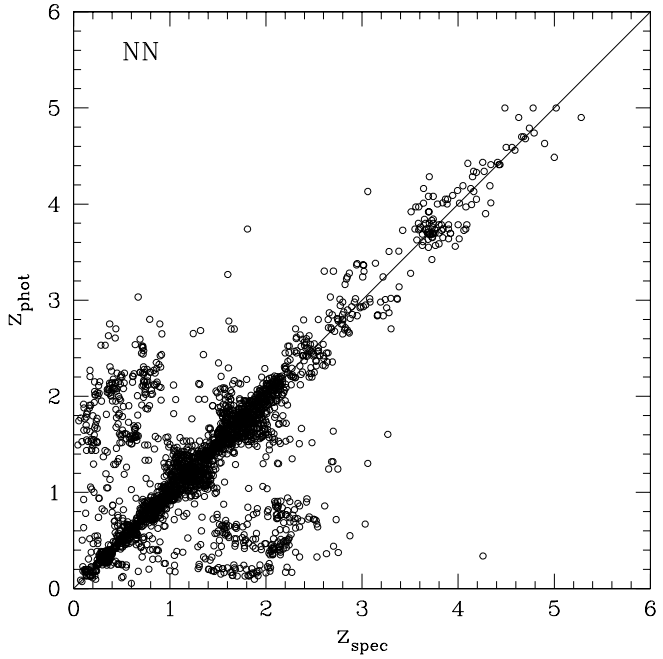

FIG. 1.—Photometric redshifts from the nearest neighbor estimator. The bulk of the objects have accurate redshifts without any systematic deviation, but the photometric degeneracy confuses quite a few estimates. The rms error is $\Delta_{\rm all} = 0.64$ for all objects and $\Delta_{0.3} = 0.116$ when excluding outliers (see text).

To determine whether the magnitude of a QSO might alleviate the degeneracy between the colors of QSOs as a function of redshift, we applied the NN method in magnitude space. We find that this degeneracy is also present when using magnitudes. One hope for removing this effect may come from extending the photometric observations to longer or shorter wavelengths, for example, using UV or IR data. Plausible reasons for the large number of redshift outliers are discussed in Richards et al. (2001b). We note that the behavior of the estimation error seems to change around $z \sim 3$, with the outliers almost completely disappearing. This is the redshift where the Lyman break passes out of the u' filter, which introduces a strong continuum feature into the quasar SED, and at higher redshifts the continuum blanketing of the Lyα forest has an even more significant effect on the observed spectrum. This attenuation enhances the quality of the redshift estimation (as is the case for high-redshift galaxies). It is also important to note that these empirical methods seriously break down outside the redshift range that the training set is in. Extrapolation is impossible, for example, to higher redshifts. The method also assumes that the set of objects used

for training is representative of the objects studied. Finally, transformation to other filter systems is extremely hard, if not impossible, without using spectral templates.

3. TEMPLATE FITTING

Template-fitting photometric redshift methods have the advantage of not requiring a training set (we assume that the physics of the QSO SEDs is fully encoded within our distribution of templates). We do, however, require a complete spectral library that has spectral templates that cover all of the spectral types of sources within our sample and which extend over the rest-frame wavelength ranges that our photometric observations occupy (for high- and low-redshift sources). In its simplest form, this template library could be the mean spectrum of a sample of SEDs. For galaxies this is not a realistic option, since the continuum shape of different spectral types of galaxies varies drastically. For quasars, there are no well-defined classes of continuum spectral features, mostly because previous observations have indicated that quasars have SEDs with similar power-law continua.

In an attempt to construct such a composite mean QSO spectrum, we consider those available within the astronomical literature. We tested the method with the Large Bright Quasar Survey composite spectrum (Francis et al. 1991). In the meantime, the SDSS composite spectrum (Vanden Berk et al. 2001) became available; in this study, we present results using this composite. In Figure 2, we plot photometric redshifts from template fitting against spectroscopic redshifts, where template colors are derived by convolving the SDSS filter functions with the SDSS composite spectrum, as well as by convolving the filter functions with our reconstructed template (Fig. 3; see below). The figure is similar to that found in Figure 1: the outliers are more stratified in the template-fitting case, and the dispersion in photometric versus spectroscopic redshift is somewhat larger than in the NN case ($\Delta_{all} = 0.89$ and $\Delta_{0.3} = 0.123$).

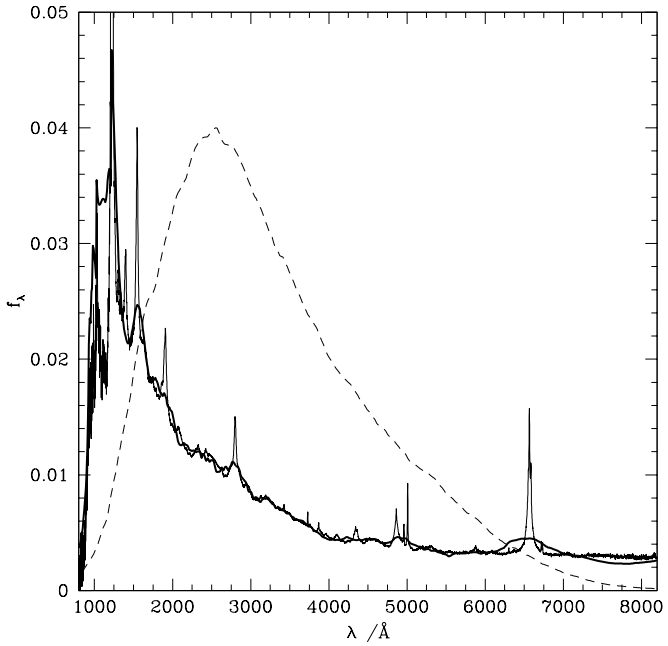

FIG. 3.—Comparison of composite SEDs. The SDSS composite QSO spectrum (*thin line*) and an average SED derived from photometry and redshifts of SDSS quasar sample are shown (*thick line*). The power-law continuum shape of the trained SED tracks the thin line nicely, and the most powerful spectral features (e.g., Lyα, C IV, Mg II) are also seen. The dashed line is the scaled constraint curve (see text for details).

Of course, direct empirical methods yield less astrophysical information—as opposed to template fitting, which provides consistent spectral type and luminosity besides redshift—and are hard to apply to different sets of observations, for example, a sample with a different filter set.

The template-fitting estimator does not seem to be superior when compared with NN estimators (see above

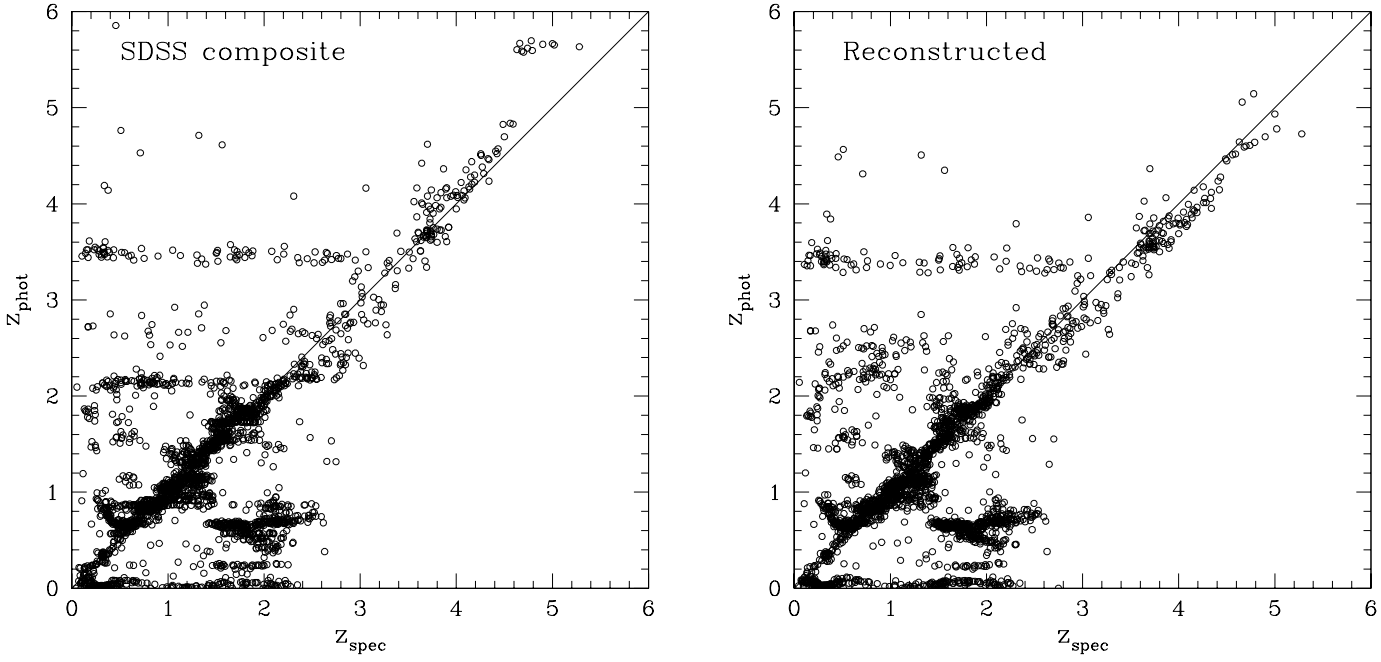

FIG. 2.—Photometric redshifts from one single SED. The SDSS composite-based estimates are seen on the left (rms errors $\Delta_{all} = 0.84$, $\Delta_{0.3} = 0.128$) and results from the reconstructed template on the right ($\Delta_{all} = 0.89$, $\Delta_{0.3} = 0.123$).

and also Richards et al. 2001b) if we use a single composite SED. In some sense, it can be considered worse; there are approximately 20% more outliers. It is likely, therefore, that the composite template assumed for the photometric redshift estimation is not representative of the full QSO sample, and that we might need to introduce additional quasar spectral types (e.g., corresponding to broad absorption lines or Fe II emission objects) to our analysis. Empirical estimators, such the NN method or that discussed in the companion paper, do not assume that QSOs at all redshifts have similar rest-frame spectra, as opposed to the template-fitting technique using a composite SED. Thus, the fact that the empirical estimators yield better redshifts than a single composite may also imply that rest-frame QSO spectra correlate with redshift. We explore these ideas in the following section.

4. TEMPLATE RECONSTRUCTION

Without having access to a complete quasar template library, we adopt a hybrid method developed by Budavári et al. (1999, 2000) and Csabai et al. (2000) to bring together the advantages of empirical and template-fitting techniques. This method requires a training set of objects with spectroscopic redshifts and multicolor photometry to establish a statistical representation of the underlying QSO SEDs. The templates are created to yield simulated colors in the best possible agreement with the photometry, and using them for photometric redshift estimation is just one application. The method is also suitable for detecting possible differences between the spectroscopic and photometric flux calibrations and to detect whether the introduction of different QSO spectral types are needed.

We propose a novel technique, similar to what is known as learning vector quantization (Kohonen 1995) in the neural network literature, that reconstructs discrete SEDs in an iterative way. The procedure, called *adaptive spectrum quantization* (ASQ), is so robust that it can develop a set of template spectra from scratch, i.e., from one single constant function. The idea here is to separate quasars into coherent classes of spectral templates based on their photometry and spectroscopic redshifts. The algorithm improves the templates step by step to increase the agreement with the photometric observations. Since the method is capable of adding and removing classes, it can be started from a single class. One loop of our iterative procedure consists of four steps:

1. All objects in the training set are categorized into classes of the most likely templates.
2. The estimated SEDs of the objects are *repaired.*
3. The reference templates are replaced with the mean of the repaired spectra.
4. The templates are dynamically duplicated or removed if statistically desired.

Let $\Psi = \{\psi_i(\lambda)\}$ represent the initial set of spectral templates, where N is the number of templates and $i = 1, \ldots, N$. Given the spectroscopic redshift and multicolor photometry of an object, the most likely template can be selected from a χ^2 optimization. Unfortunately, in general, even the best-fitting template $\psi_k(\lambda)$ is not perfect, in the sense that it cannot reproduce the measured colors if convolved with the throughput of the filters and instrument response. So, the $\psi_k(\lambda)$ spectrum must be "repaired" according to the observational constraints (i.e., the multicolor photometry) as described in Budavári et al. (2000). The adjusted spectrum

$$\psi'_k(\lambda) = \psi_k(\lambda) + \delta\psi_k(\lambda) \quad (1)$$

will be in better agreement with the measurements, but it may be affected by some irregularity of the object. So, instead of replacing the original $\psi_k(\lambda)$ with $\psi'_k(\lambda)$ or adding the repaired SED to the original Ψ set, one can collect all of the adjusted spectra assigned to the kth template and replace $\psi_k(\lambda)$ with the average of these repaired template spectra. A similar reparation can be applied to all templates used in the analysis:

$$\psi_k(\lambda) \rightarrow \langle\psi'_k(\lambda)\rangle = \psi_k(\lambda) + \langle\delta\psi_k(\lambda)\rangle\ . \quad (2)$$

Introducing additional reference templates requires more deliberation. One can think of several conditions in which a class of objects should be split into two (e.g., based on the number of assigned objects), but when the observed training set is limited in size, the choice of this condition may be crucial. When the iteration converges with a fixed number of templates (relaxation), a new reference spectral template can be added to the set if deemed necessary because of a large intrinsic scatter around the mean repaired spectrum $\langle\psi'_k(\lambda)\rangle$ or because of an excess of member objects within a particular spectral class. To split a class, one can make a copy of the corresponding template and perturb it slightly by shifting its spectral shape toward that of the modulated spectrum of a random member within its spectral class. A more elaborate discussion of the method will be described elsewhere (Budavári et al. 2001).

We tested the above method for galaxies—where the photometric redshift estimation techniques are well established—and it provides photometric redshifts for galaxies that are in good agreement with those estimates based on eigentemplates. Quasar redshifts can also be predicted with accuracy. Starting from a featureless, constant SED,

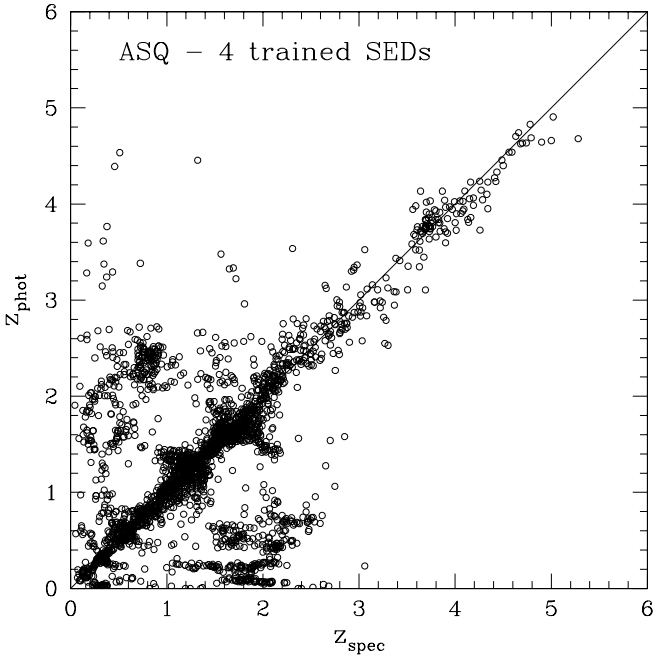

FIG. 4.—The ASQ training yields templates that provide reliable redshift estimates. Here we plot the results using four reconstructed templates ($\Delta_{all} = 0.77$, $\Delta_{0.3} = 0.120$). The degeneracy is clearly present in the plot just as in the NN estimates.

we develop an average template that is very similar to the empirical SDSS composite spectrum (Vanden Berk et al. 2001), as seen in Figure 3. Our reconstructed template (derived purely from the photometric data) tracks the power-law continuum of the high-resolution spectrum to a remarkable degree of accuracy. The good correspondence between the continuum shapes shows not only the strength of the reconstruction algorithm, but also that the efforts the SDSS team made for the correct spectrophotometric calibration were fruitful.

The strongest spectral lines, such as Lyα (at 1216 Å, merged with N v at 1240 Å), C IV (1548 Å), and Mg II (2798 Å) are clearly visible. The Hβ (4861 Å) and the [O III] (4959, 5007 Å) lines are merged into a single feature, and the blurred Hα line (at 6563 Å) is also present. The sharpness of these lines is quite surprising, given that only the broadband photometric information (a typical filter width is 1500 Å) and the value of the spectroscopic redshift were used as input parameters in the reconstruction algorithm.

For a given redshift, the photometric observation gives constraints on the possible underlying SED, since we expect to get back the measured photometric values by redshifting the SED and convolving it with the filter response function. This constraint obviously depends on the photometric system and, also, the redshift of the object as the rest-frame spectrum is sampled at different wavelengths. Adding up the filter curves, shifted to the rest frame, for all objects provides a smooth curve as a function of wavelength (Fig. 3, *dashed line*) that is proportional to the amount of information available for the determination of the shape of the composite. This is similar to what is typically shown for composites derived from spectrophotometric observations, but computed from the shape of the photometric transmission curves. The more objects at different redshifts we use to constrain the spectrum, the finer the details that could be recovered, for example, the width of the reconstructed line can be smaller. An additional effect is that short-wavelength lines are constrained by higher redshift objects so that when they are transformed back to the rest frame their line widths become narrower. In Figure 3, one can see that reconstructed lines at shorter wavelength are narrower than those at longer wavelength.

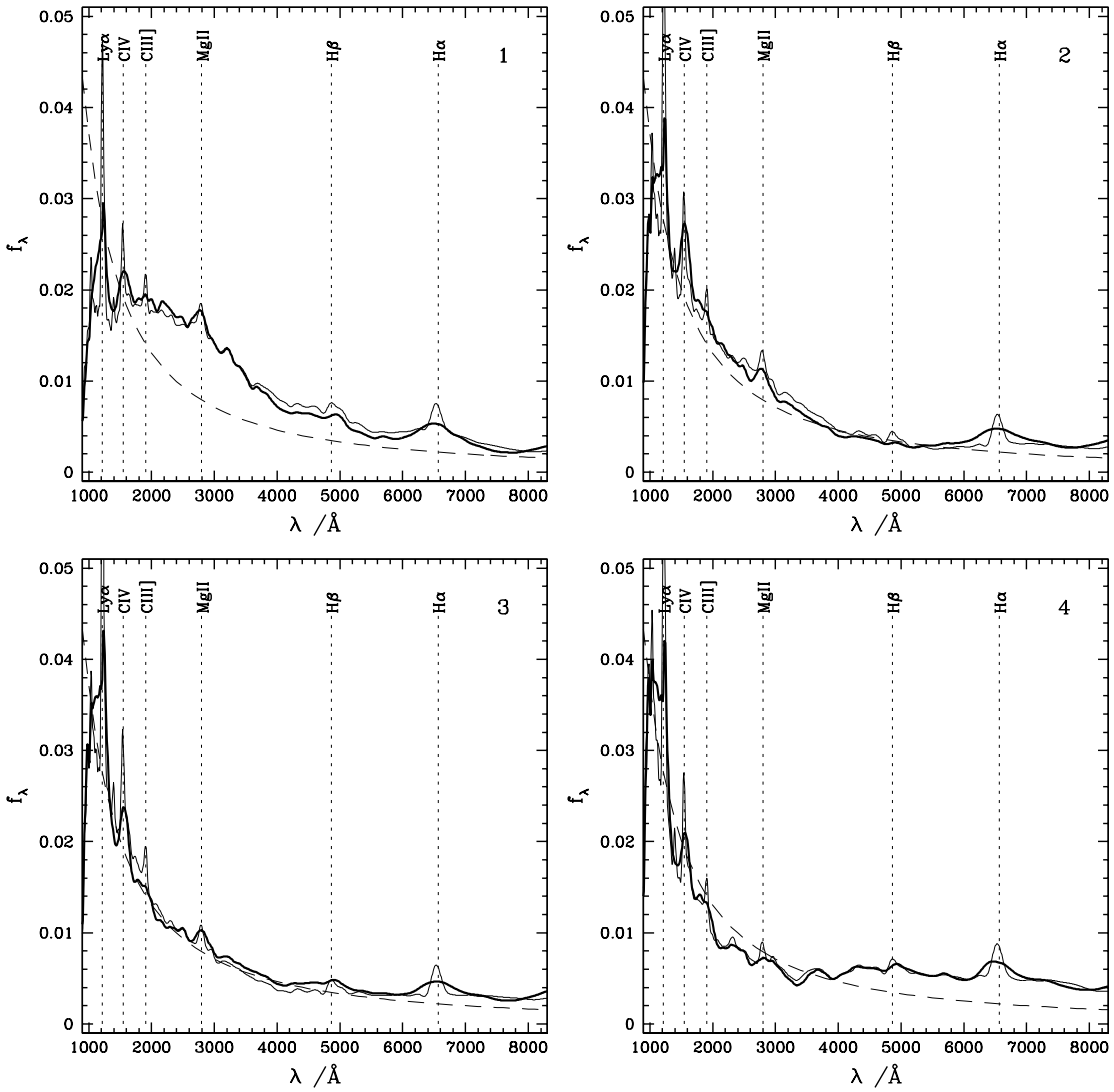

FIG. 5.—Four templates reconstructed from a constant initial SED (*thick line*) and from the composite SDSS SED (*thin line*). Despite the different initialization, the reconstruction algorithm converges to similar sets of templates. The strongest known quasar lines (*dotted vertical lines*) appear in the reconstructed spectra. The same power-law curve ($f_\lambda \propto \lambda^{-1.5}$) is plotted in each panel (*dashed line*) to guide the eye to see the differences.

The difference between the composite and reconstructed spectra below ~ 1200 Å is due to the fact that the composite spectrum is not corrected for intergalactic attenuation. Comparing the photometric redshift estimations with the composite and the reconstructed spectrum in Figure 2, we find little difference except at high redshifts, which is due to the effect of a different handling of the intergalactic attenuation.

5. QSO SPECTRAL TYPES

Now we arrive at the last step of the ASQ algorithm, where we consider whether there is evidence for more than just one spectral template. The optimal number of templates is partly determined by the size of the training set. If we divide a sample into too many subsets, there will not be sufficient objects within any one subset to constrain the reconstructed SEDs. On the other hand, the more spectral types we allow, the closer our resulting template library will be to the SEDs of the quasars for which we wish to predict a redshift. Our experiments show that, for the current training set, a QSO template library with four templates is the optimal choice. Introducing further templates does not improve the agreement of synthetic and measured colors significantly (χ^2 decreases less than 10%), and the photometric redshifts do not become significantly better.

Compared with Figure 2, in Figure 4 the number of outliers at $z_{\rm phot} \approx 0$ and $z_{\rm phot} \approx 3.5$ is greatly reduced, and the redshift estimation improves by increasing the number of templates from one to four. Using the four ASQ templates yields $\Delta_{\rm all} = 0.77$ and $\Delta_{0.3} = 0.120$, which is still above the results of empirical methods. The algorithm can be started from a single constant function or from the empirical composite SED. Both starting points lead to templates with approximately the same continuum features (see Fig. 5). The only difference is that the emission lines remain sharper in the latter case. The match between the two sets is even more interesting if we remember that the two template libraries were created using significantly different initial spectra. This shows that the classification of quasars into these spectral types is robust.

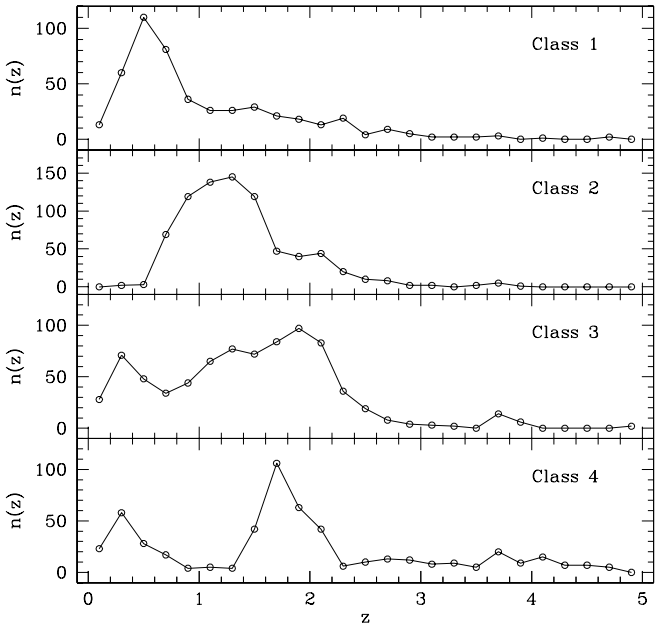

FIG. 6.—Redshift distribution of objects in the four ASQ template classes.

During the photometric redshift estimation with the template-fitting method, each object is assigned uniquely to a spectral template derived earlier by the ASQ algorithm. To test whether there is any correlation between the redshift of the objects and their spectral class, we plot the redshift distributions of the four ASQ classes in Figure 6. We see a relation between type and redshift, which may have several causes. Among other possibilities, it could be related to the different mean absolute luminosity of the classes, different mixture of broad absorption line (BAL) and other quasar types, or sampling issues, or it may be a sign of QSO spectral evolution. It is also important to note that at different redshifts the photometric passbands sample different portions of the templates, and this can be accounted for by having low-redshift objects in classes 3 and 4, where the highest redshift quasars are. Detailed analysis of this is subject of a future work.

We created composites for each ASQ spectral class using SDSS spectrophotometric observations. In Figure 7, we compare the resulting composites with the four ASQ templates. The order of the templates here and also in Figure 5 corresponds to the increasing mean redshift value of quasars in the classes. Comparing the continuum shape of the ASQ SEDs (the same power law $f_\lambda \propto \lambda^{-1.5}$ is plotted with dashed lines to aid in comparison) suggests a trend with redshift.

Aside from the apparent redshift dependence of each class, there are at least two other distinct types of quasars that seem to be at least partially segregated into the different classes: BAL quasars and low-redshift, low-luminosity quasars with host galaxy contamination.

Broad absorption features are evident in the class 1 composite spectrum and also, to a lesser extent, in the class 2 composite spectrum. The BAL QSOs in the quasar sample have been identified, and the majority (71%) fall into either class 1 or class 2. The fraction of quasars that have BAL features in their spectra for each class (1, 2, 3, 4) is 11%, 6%, 1.9% and 3%, respectively. The colors of BAL QSOs tend to be redder than non-BAL quasars (Menou et al. 2001; Sprayberry & Foltz 1992; Brotherton et al. 2001; Weymann et al. 1991), which probably accounts for their unequal separation here into the different classes.

The evidence for host galaxy contamination comes from the presence of stellar absorption lines in the composite spectra. For example, the Ca II H and K lines (Fig. 7, *insets*) are clearly present in the class 4 composite and weakly present in the class 3 composite. There is no evidence for them in either the class 1 or class 2 composite. Host galaxy contamination is likely to be a significant contributor to the reddening of quasars and other active galactic nuclei at wavelengths beyond about 5000 Å (Vanden Berk et al. 2001). For redshifts less than 1, the quasars in classes 3 and 4 are significantly less luminous than their counterparts in classes 1 and 2, and the quasars with the lowest redshifts tend to belong to classes 3 and 4. This points to host galaxy contamination in the spectral light of the low-redshift quasars in classes 3 and 4 and is likely the cause of their being preferentially separated into those classes.

Now we return to the problem of the photometric degeneracy (i.e., quasars with the same photometric colors can have significantly different redshifts). Signs of the photometric degeneracy of quasars are visible in the photometric

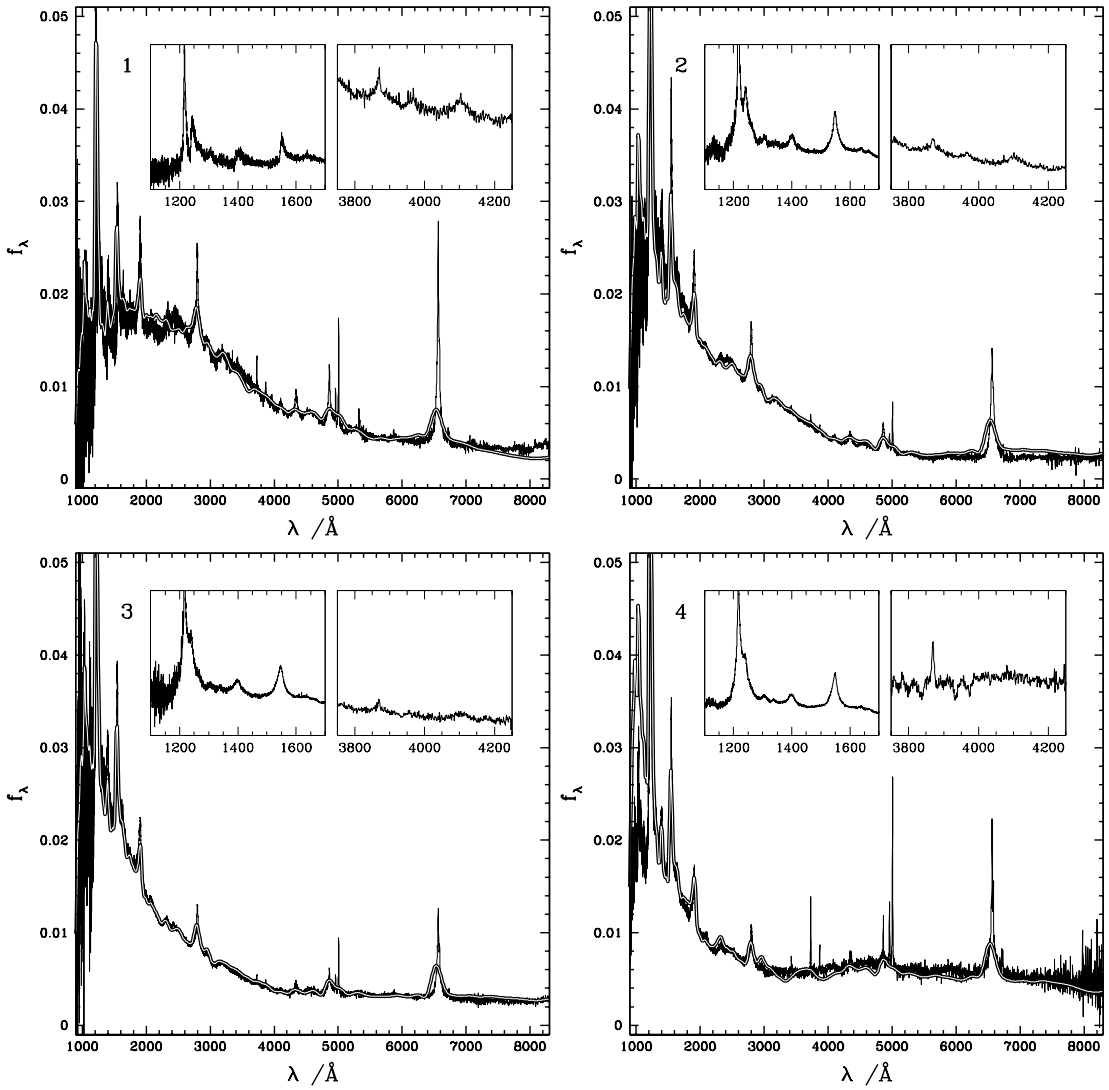

FIG. 7.—Comparison of the ASQ templates (*white lines*) with composites created for the corresponding classes using SDSS spectra (*thin lines*). See discussion in text. The insets zoom in on the Lyα to C IV region and the Ca II H and K lines.

redshift prediction process. The χ^2, roughly the weighted distance of the observed and estimated flux as a function of redshift z, has multiple minima as a result of the degeneracy in photometry. Two outliers from the photometric redshift relation are shown in Figure 8. Both have significant minima at the spectroscopic redshift.

In Figure 9, we plot what one would obtain if one could break this degeneracy by using further observations (e.g., UV) or through the use of priors that can crudely constrain the possible redshift of an object. The photometric redshift estimation algorithm can find all of the local minima of the χ^2 curve, and hence we could have more redshift candidates for an object to choose from. Given the best three local minima for each object, we selected the one having the closest value to the spectroscopic redshift; this somewhat idealized case is plotted in the figure.

6. CONCLUSIONS

The success of our attempt to determine redshifts of quasars from SDSS multicolor photometry, along with the Richards et al. (2001b) companion paper, is encouraging. We can obtain redshift estimates for 70% of the QSOs with accuracy $\Delta z_{\rm rms} \approx 0.1$. The method discussed in this paper is not only able to obtain redshifts, but also to estimate luminosities and SEDs for QSO candidates without spectroscopy. Photometric redshifts are ideally suited for many astrophysical and astronomical studies, such as the luminosity function and gravitational lensing.

Our spectral template reconstruction appears to be very robust; the algorithm yields similar sets of spectral templates whether started from a constant function or from a composite spectrum of spectrophotometric measurements. Individual spectral lines have significant effects on quasar colors. These narrow features (compared with photometric bands) contribute to photometric observations sufficiently that they can be recovered by our reconstruction technique. It is quite reassuring that the spectral templates reconstructed from the very low resolution photometric data show the well-known emission lines of quasars. This high resolution is possible because of the large number of SDSS objects used in the template reconstruction process. Detailed analysis of the templates can yield interesting information on quasar spectral features. We found that the use of multiple templates in the redshift estimation

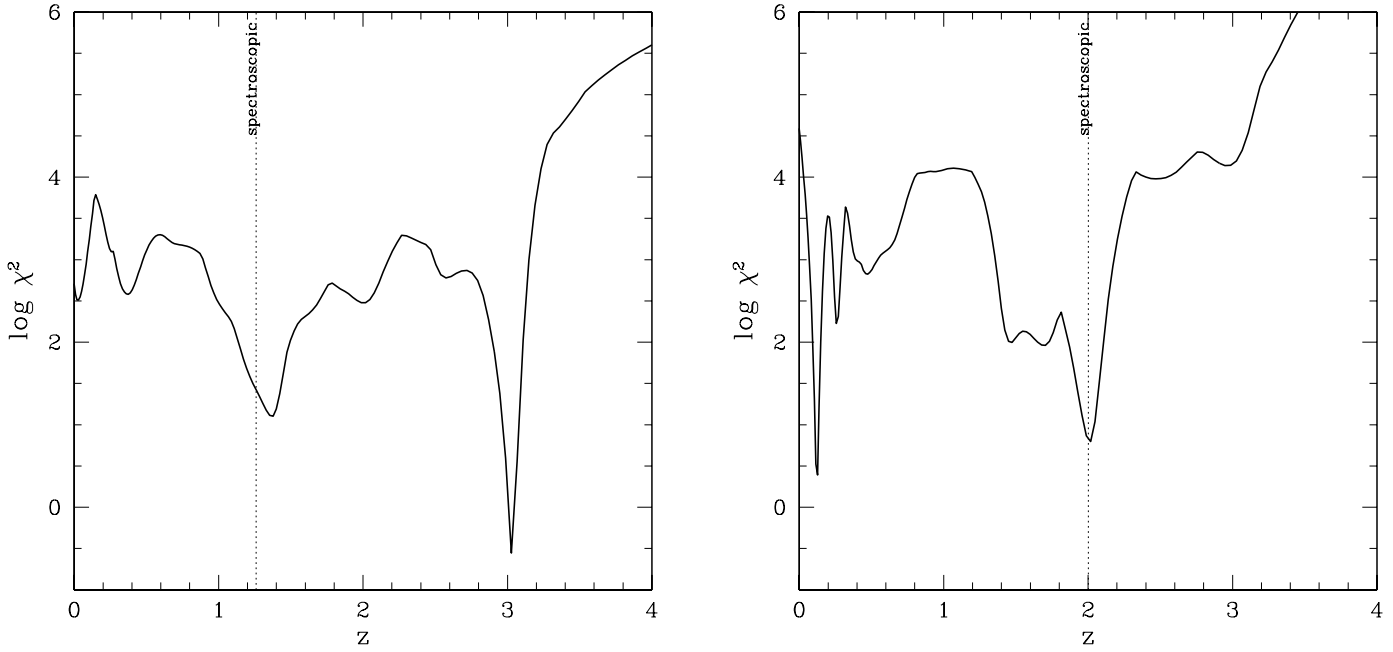

FIG. 8.—The χ^2 curve describes the agreement of the most likely template and the photometric measurements at an assumed redshift. The figure illustrates the degeneracy in photometry by plotting χ^2 vs. z, where more minima are present. For outliers the algorithm gives a false redshift value at the global minimum, but there is also a significant local minimum at the true spectroscopic redshift (*vertical dotted line*).

improves the photometric redshift relation and that the resulting spectral types correlate with redshift.

The significance of the spectral lines in broadband photometry also means that potential errors in the boundary of the assumed passbands may result in more serious problems in the photometric redshift estimation than is the case for galaxies, where broadband spectral features dominate. If the edge of a photometric filter were off just by the width of a spectral line, then the colors, and hence the estimated photometric redshifts, would change substantially. The fact that the reconstructed spectral lines match well the location of the actual features is an indication that the filter curves are indeed accurately calibrated. For surveys with both photometric and spectroscopic observations (such as SDSS), our method is well suited to cross-check calibration (e.g., to photometrically calibrate spectroscopic observations) or may be used to detect systematic drifts in annual, seasonal, etc., data sets.

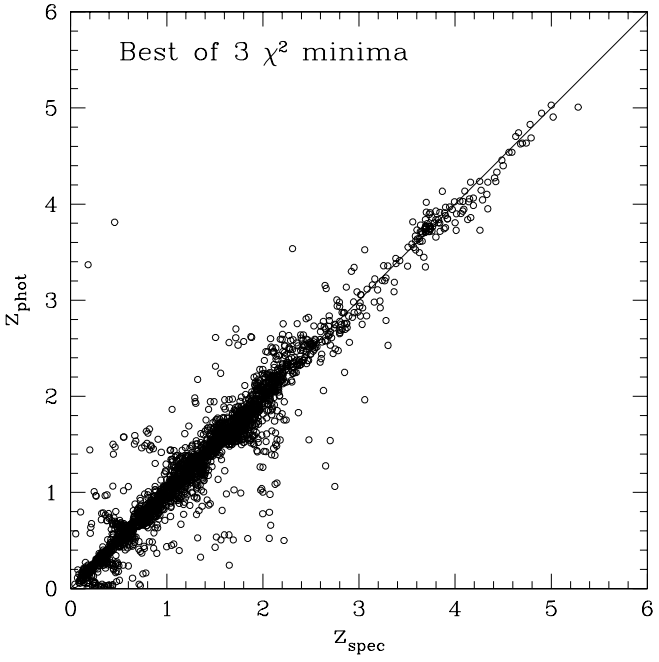

FIG. 9.—The same ASQ templates (see Fig. 4) could provide much better redshift estimates if the photometric degeneracy could be broken, e.g., by applying some prior. The figure shows a somewhat idealized case: we use the closest z_{phot} value to z_{spec} from the three smallest χ^2 minima ($\Delta_{all} = 0.238$ and $\Delta_{0.3} = 0.104$).

The template reconstruction method can be extended to derive photometrically calibrated SEDs for objects fainter than the limit at which spectrophotometric observations are available. Rough photometric redshift estimates can be also used in the algorithm as long as there are no systematic deviations, only scatter around the true value. An iterative process could be developed to fine-tune the templates. It could be especially useful for faint high-redshift QSOs, where spectroscopy is not available, but the Lyα attenuation could give a fairly good redshift estimate.

The Sloan Digital Sky Survey is a joint project of the University of Chicago, Fermilab, the Institute for Advanced Study, the Japan Participation Group, Johns Hopkins University, the Max-Planck-Institut für Astronomie, the Max-Planck-Institut für Astrophysik, New Mexico State University, Princeton University, the US Naval Observatory, and the University of Washington. Apache Point Observatory, site of the SDSS telescopes, is operated by the Astrophysical Research Consortium. Funding for the project has been provided by the Alfred P. Sloan Foundation, the SDSS member institutions, the National Aeronautics and Space Administration, the National Science Foundation, the US Department of Energy, the Japanese Monbukagakusho, and the Max Planck Society. The SDSS Web site is http://www.sdss.org/. I. C. and T. B. acknow-

ledge partial support from Hungarian Academy of Sciences–NSF grant 124 and Hungarian National Scientific Research Foundation grant T030836. A. S. acknowledges support from the NSF (AST 98-02980) and the NASA Long Term Space Astrophysics program (NAG 5-3503). A. J. C. acknowledges partial support from NSF grants AST 00-96060 and AST 99-84924 and NASA LTSA grant NAG 5-8546.

REFERENCES

Benítez, N. 2000, ApJ, 536, 571
Brotherton, M. S., Tran, H. D., Becker, R. H., Gregg, M. D., Laurent-Muehleisen, S. A., & White, R. L. 2001, ApJ, 546, 775
Brunner, R. J., Connolly, A. J., & Szalay, A. S. 1999, ApJ, 516, 563
Budavári, T., Szalay, A. S., Connolly, A. J., Csabai, I., & Dickinson, M. E. 1999, in ASP Conf. Ser. 191, Photometric Redshifts and High Redshift Galaxies, ed. R. J. Weymann, L. J. Storrie-Lombardi, M. Sawicki, & R. Brunner (San Francisco: ASP), 19
———. 2000, AJ, 120, 1588
Budavári, T., et al. 2001, in preparation
Connolly, A. J., Budavári, T., Szalay, A. S., Csabai, I., & Brunner, R. J. 1999, in ASP Conf. Ser. 191, Photometric Redshifts and High Redshift Galaxies, ed. R. J. Weymann, L. J. Storrie-Lombardi, M. Sawicki, & R. Brunner (San Francisco: ASP), 13
Connolly, A. J., Csabai, I., Szalay, A. S., Koo, D. C., Kron, R. G., & Munn, J. A. 1995, AJ, 110, 2655
Csabai, I., Connolly, A. J., Szalay, A. S., & Budavári, T. 2000, AJ, 119, 69
Fernández-Soto, A., Lanzetta, K. M., & Yahil, A. 1999, ApJ, 513, 34
Francis, P. J., Hewett, P. C., Foltz, C. B., Chaffee, F. H., Weymann, R. J., & Morris, S. L. 1991, ApJ, 373, 465
Fukugita, M., Ichikawa, T., Gunn, J. E., Doi, M., Shimasaku, K., & Schneider, D. P. 1996, AJ, 111, 1748
Gwyn, S. D. J., & Hartwick, F. D. A. 1996, ApJ, 468, L77
Kohonen, T. 1995, Self-organizing Maps (Berlin: Springer)
Koo, D. C. 1985, AJ, 90, 418
Lupton, R. H., et al. 2001, in preparation
Madau, P. 1995, ApJ, 441, 18
Menou, K., et al. 2001, BAAS, 198, No. 78.04
Richards, G. T., et al. 2001a, AJ, 121, 2308
———. 2001b, AJ 122, 1151
Sawicki, M. J., Lin, H., & Yee, H. K. C. 1997, AJ, 113, 1
Sprayberry, D., & Foltz, C. B. 1992, ApJ, 390, 39
York, D. G., et al. 2000, AJ, 120, 1579
Vanden Berk, D. E., et al. 2001, AJ, submitted
Wang, Y., Bahcall, N., & Turner, E. L. 1998, AJ, 116, 2081
Weymann, R. J., Morris, S. L., Foltz, C. B., & Hewett, P. C. 1991, ApJ, 23, 53
Weymann, R. J., Storrie-Lombardi, L. J., Sawicki, M., & Brunner, R., eds. 1999, ASP Conf. Ser. 191, Photometric Redshifts and High-Redshift Galaxies (San Francisco: ASP)